DOCUMENT RESUME

ED 260 678 HE 018 807

AUTHOR Loacker, Georgine; And Others
TITLE Assessment in Higher Education: To Serve the Learner.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Oct 85
NOTE 43p.; Paper presented at the National Conference on Assessment in Higher Education, sponsored by the American Association for Higher Education, (Columbia, SC, October 13-15, 1985).
PUB TYPE Guides - Non-Classroom Use (055) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Academic Ability; Academic Standards; College Faculty; *College Students; *Evaluation Criteria; *Feedback; Guidelines; Higher Education; Intellectual Development; *Outcomes of Education; *Self Evaluation (Individuals); *Student Evaluation
IDENTIFIERS Alverno College WI

ABSTRACT

The idea of assessment as learning is discussed, along with the practical history of assessment, and how student assessments should be conducted. Background information covers the use of assessment in business and government, as well as in education and psychology. It is proposed that assessment include: multidimensional sampling of students' abilities in action; observation and judgment of those samples on the basis of explicit criteria; and structured feedback, administered sequentially in relation to a learner's development. Each of these elements in turn must contribute to the growth of the students' ability to self-assess. Issues relevant to beginning, intermediate, and advanced learners are identified, and research on criteria, feedback, and self-assessment is reviewed. To help faculty design individual classroom assessments, consideration is given to: determining a specific ability or expected outcome, identifying component abilities, selecting or designing a stimulus and context, attending to developmental levels, developing criteria, providing for self-assessment, and judging performance and giving feedback. Included are guidelines for judging whether criteria are effective and an example of extending assessment beyond the classroom to a wider curriculum context. (SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED260678

ASSESSMENT IN HIGHER EDUCATION: TO SERVE THE LEARNER

Georgine Loacker
Lucy Cromwell
Kathleen O'Brien

Alverno College Milwaukee, Wisconsin

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

HE 018 807

A paper prepared for the American Association for Higher Education, under contract with the National Institute of Education, for the Conference on Assessment in Higher Education at the University of South Carolina, Columbia, October 13-15, 1985.

Assessment seems to be loitering expectantly in the corridors of higher education, thereby reinforcing the hope that it will soon enter the classroom to serve the learner. National higher educational reports today encourage assessment. Administrators call for it. Researchers see it as a potential instrument for prediction and evaluation. Legislators look to it for assurance of accountability. But many of these intents overlook the power of assessment for teaching and learning. So that teachers might take a more serious look at assessment, we propose in this paper to set it at the heart of the learning dynamic and clarify it as a major strategy to be used by the teacher and placed in the hands of the learner.

The first use of the word assessment did not emerge from classroom or campus, but the meaning of the word began with an idea important to educators -- that of sitting down beside or together from late Latin ad+sedere. In the seventeenth century an assessor was "one who sits beside" or "who shares another's position." Early uses of the word focused primarily on determining the worth or value of something in monetary terms, but consistently underlying it was the element of skilled or expert judgment made on the basis of careful observation. Thus it seems to be a word destined for tongues of educators -- whether humanists or scientists.

PLAN OF THIS PAPER

In probing assessment's use in education, we begin this paper by developing a series of assumptions as direction-setters for what follows. After some definitions and distinctions, we continue via three connected routes, all of which lead to the central idea of assessment as learning. The first route is historical: it shows how assessment emerges from business, psychology, and education with implications for what it can do for the learner. The second route is conceptual: it analyzes elements that must be present, individually and in relation to each other, to constitute assessment. The third route brings these themes home: it shows how faculty can, individually and collaboratively, create an assessment process that serves their students.

Except for historical and research references, the propositions in this paper are based on the cumulative experience of the authors with students. In collaboration with our college's faculty as a whole, we continue to test our theory in the classroom and through ongoing institutional research.

DEFINITIONS AND ASSUMPTIONS

Assessment, as we use it throughout these pages, is a multidimensional process of judging the individual in action. Embedded in this definition are assumptions about learning that emphasize active development of the learner.

Assumptions. One assumption is that learning involves making an action out of knowledge -- using knowledge to think, judge, decide, discover, interact, and create. We contend that acquiring or storing knowledge is not enough. Unless one carries knowledge into acts of application, generalization, and experimentation, one's learning is incomplete.

Another assumption is that an educator's best means of judging how well a learner has developed expected abilities is to look at corresponding behavior -- thinking behavior, writing behavior, inquiry behavior, or appreciating behavior, for instance. We presuppose a link between behavior and cognitive and affective processes. Because human behavior is purposeful, educators can find out more about a learner's problem solving ability by observing that person's solving process than by confirming a "correct" solution he or she has selected from a set of alternatives.

A third assumption is that learning increases, even in its serendipitous aspects, when learners know what they are setting out to learn, understand what standards they must meet, and have a way of seeing what they have learned. When students of science, for example, know that they will have to go beyond reading their text, listening to their teacher, and replicating lab experiments, that they will have to raise their own questions and test their own explanations, they are more apt to learn to do all of the above more meaningfully and effectively. Out of that success they then develop

assurance that enables them to recognize unsought-for insights when they come upon them.

We contend that such awareness of expectations and standards enhances learning because it places in a person's hands the means of collaborating in his or her own learning and gradually taking control of one's own learning process. Within that context, learners recognize that their question, "How am I doing?" is taken seriously. They begin to see an important implication of that question: that further learning builds on and develops out of where each learner is at any given point. Therefore, that question becomes the occasion for doing better when everyone responsible for the learning -- teacher as well as student -- gets as complete an answer as possible. Assessment aims for such an answer.

What does it mean to aim at an increasingly complete answer to the question of how a person is doing? One can get some insight into the question by looking at what testing traditionally tells us about someone, in contrast to what assessment does.

Assessment in Contrast to Testing and Measurement. Testing can tell us how much and what kind of knowledge someone has. Assessment gives us a basis for inferring what that person can do with that knowledge. Testing carefully limits to a set of written or marked answers what we can know about a person. Assessment aims to elicit whatever a person can do to show the nature, extent, and quality of his or her ability.

Testing focuses on measurement. It answers the question "How am I doing?" with a quantitative response that says, "You did a certain percent of what was asked on a given occasion" or "You did as well as a certain percent of all those who tried or might try to do the same." Assessment answers the question with a descriptive account of precisely what the individual person has done on a given occasion. By judging a person's performance against criteria, assessment aims to give it a meaning out of which he or she can build future performance.

Assessment and Evaluation. Emphasis on the progress of the individual learner also distinguishes assessment from program evaluation. Assessment looks for distinguishing elements in a person's performance. Evaluation looks for elements that can be combined and compared in order to draw conclusions about groups of students, with a view to making judgments about the general direction of a course, program, or curriculum. Testing, measurement, and evaluation all rely on consistency of context. Assessment relies on varying of contexts to assure the eliciting of as much complexity of a person's ability as possible.

Our definition of assessment is shaped by its power to serve the learner: it means eliciting samples of varied expressions of an ability, judging those samples against identified criteria, and providing as full a picture as possible of that ability as possessed by that learner. Assessment as learning weaves together several strands of a long history of meaning that have developed separately.

I. HISTORY
Assessment in Business and Government

The practical history of assessment in business and government is essentially the history of the Assessment Center Method. And it is -- at least until recently -- the history of improved selection and screening rather than of development and learning. In the 1930's, when it began in England and Germany, assessment provided a new, behaviorally oriented means of selecting military officers. In the 1940's, with researchers from the Harvard Psychological Clinic adapting and further developing assessment, the United States Office of Strategic Services used it to select American intelligence agents. In the 1950's, led by AT&T, business and non-military government departments contributed to the extensive growth of assessment centers by using them to select managers. More recently, business and government have begun to show interest in using the assessment center method for development.

The spread of assessment centers has established, as part of the connotation of assessment, several concepts from which education can benefit. Assessment, as it

characterizes the assessment center method, connotes using behavioral descriptors to develop a richer picture of an individual's ability, using multiple techniques for performance assessment, and refining expert judgment through articulation of more explicit evidence.

Assessment in Education and Psychology

When educators speak of assessment, they do not typically use the term to mean a focus on the individual learner, the employment of multiple instruments, or an attempt to gauge complex behavior. An ERIC search of the higher-education literature published between 1972 and 1985 yields 223 articles that combine the key word, assessment, with one or more of the following terms: feedback, student achievement, student evaluation, and outcomes of education. A review of these articles indicates that the term assessment is used in broad, varying senses. Its most frequent use is as a synonym for program evaluation. However, in all of the above contexts, as well as in the context of clinical psychology, the word assessment emerges often as a contrast to testing and thus creates, as part of its connotation, a concern with broader educational outcomes than knowledge.

A concrete example of such a use of the word since the 1970's is the National Assessment of Educational Progress project. Although this program is interested in group patterns rather than individual development, its concern was to create objectives that incorporated skills and attitudes as well as knowledge and to make sure that the instruments used corresponded with clear objectives.

A related use of the word assessment is by the Council for the Advancement of Experiential Learning (CAEL). Since the organization began in the 1970's, members' use of the word assessment for adult learners has promoted a connection between the word and the wide range of abilities that adults might develop experientially.

Psychologists, both educational and industrial, influenced by developments in personality theory, have also tended, since the 1940's and 50's, to use the term

assessment in relation to broad and multiple abilities. They have added to the connotation of the word the concept of abilities, taken not as static traits but as process, and thus changeable and directable.

In Great Britain, the use of the word assessment also increased among educators, especially since the 1960's when the idea of continuous assessment was introduced in response to criticisms of final examinations.

Thus from multiple sources, this strong association of the word with multiple performances and a breadth of abilities has clung to the connotations of the word even as the denotation of the word has become more general and diffuse.

There are several indications that the educational world is now calling for the assessment concept as we define it in this paper. National reports are asking educators for accountability in making the learner their focus in designing learning experiences. The phenomenon of increasing enrollment of adult learners is reminding educators that learning is developmental and continues throughout the life span. It also suggests that one cannot evaluate experience as prior learning unless one defines learning in terms of developed abilities or significant expected outcomes. Emphasis on experiential learning -- doing what one knows -- has surfaced as an important component of the learning cycle of an individual.

Finally, current questions about the usefulness of intelligence measures and of standardized tests such as the SAT also focus attention on the need to develop other approaches to assess a learner's ability and potential to learn.

As we have shown, specific connotations of assessment keep emerging. The result is a certain vagueness in any general use; of the word it becomes necessary to redefine the word for each context. The definition we set forth below provides a focus for our use of it in a teaching and learning context.

II. CONCEPTUAL ELEMENTS OF ASSESSMENT

Our definition of assessment is one with which teachers can feel at home. We say that it is a multidimensional process of judging the individual in action. Thus, it focuses on an important element of the best practice of the committed teacher: the close observation and careful judgment that uncover the occurrence of learning in an individual.

Every teacher has had his or her own experience of hearing some version of the young Helen Keller's cry of "Wa-ter," the experience of discovering a student's sudden illumination or success. And having once heard it, who does not wish to find a way of making it more frequent, more developmental, and more characteristic of every student? There are many things in a classroom to keep teachers from frequently discovering individual learning especially among less assertive and articulate students. Even for the brightest students, teachers need to find ways to assist in building on and expanding moments of learning rather than merely rewarding them.

Assessment becomes a meaningful way to expand learning when one defines it to include a set of key elements that make it a learning experience. It provides a way of re-focusing education on individual learners instead of putting on a wide lens that lumps them all together into an indistinguishable mass from which we can infer only general patterns. Since students come to us in groups, in classes and courses, the idea of using assessment as a camera that takes individual portraits instead of group pictures needs careful delineating. Essential is the requirement that a picture of a student's abilities be made visible to everyone responsible for the student's learning -- including the student him or herself. Crucial is the concept that the picture become dynamic, cumulative, and composite.

To create such a picture, assessment needs to be defined to include multidimensional sampling of students' abilities in action, observation and judgment of those samples on the basis of explicit criteria, and structured feedback -- administered sequentially in relation to a learner's development. Each of these

elements in turn must contribute to the growth of the students' ability to self-assess.

Sampling Student Performance

Looking at a student in action is necessary because it gets us as close to an individual's ability as we can get. Because we cannot observe all of a person's expressions of a given ability, we take intermittent samples. Given the complexity of the human person, there will always be a distance between behavioral data and knowing an ability fully. Even a very precise image of exactly how a detective has gone about solving a single mystery gives a very limited view of his or her full detecting powers. Nonetheless, sampling gives a teacher a start toward developing a picture of an ability in operation.

Without a behavioral sample, teachers can look at a set of correctly selected answers and say that a person was able to recognize given facts. They can look at a description of what a person says that he or she knows about something and would do with the knowledge. But those teachers can only assume and hope that the knowledge can translate into effective action when it needs to.

With a behavioral sample, teachers can at least see that a person did or did not do something in a given context. Therefore they can say that someone can do it -- at least in such a context -- whether or not he or she will do it again. For the persons assessed, sampling holds up an individual picture that enables them to look from the outside at their own ability in action, to supplement their inside view.

To assess, therefore, requires that we sample students' behavior. We need to sample their writing to judge whether they can write. We need to sample their synthesizing to judge whether and how they have put together the facts they have learned. We need to sample students' work in groups to judge whether they can think and work collaboratively.

Multidimensionality. The complexity of the human person, indeed of an individual human ability, suggests that sampling for assessment cannot be single or final. In order to bring to light enough dimensions for fair judgment, the sampling needs to be multiple and varied.

Multidimensionality provides a means of addressing some of the questions that single sampling raises: How do we know a sample is representative? Is the person having an unusually good or bad day? Can the person repeat the performance under different circumstances? Will he or she? The only way we can begin to form an answer is to gather enough samples to enable a pattern to emerge. Perhaps it turns out that an unusually good day -- or a bad one -- seems to be representative, or more likely, that either or both are occasional occurrences. In either case, the important thing is to be as precise as possible in discerning the elements that constitute each performance so that the assessee's knowledge/experience of them can reinforce and refine the effective and transform the ineffective.

How varied need the samples be to suggest the complexity of an ability? Given the mystery of the human person -- which can and should be only partly visible to another -- we have no way of being specific about the variety of examples needed to demonstrate an ability. Such things as a written or oral mode, a static or dynamic object of analysis, a solitary or collaborative responsibility for accomplishing a task -- all evoke different dimensions of an ability. Being able to analyze written data at one's desk, for example, does not mean being able to analyze data as it occurs before one's eyes in a group situation. Nor can good writers always organize their thoughts as well when speaking.

The reasonable response to the question of varying contexts, therefore, seems to be to vary them according to the ordinary shifts of life situations -- such things as the purpose, the nature and number of people involved, the amount of information available, the potential and limitation of form. If careful feedback is provided, each shift in context can assist learners to refine their understanding of an ability

and how they exercise it. The effect of varying context is twofold: it reinforces the general core skills involved and it reveals unique skills elicited by each situation.

The success of the "writing across the curriculum" movement dramatizes the new understanding of secondary educators that effective writing as a life ability needs to be practiced and assessed in a variety of disciplines -- in fact in every discipline studied. If writing across the curriculum makes good academic sense, why not the assessing of every ability across the curriculum?

How many samples are necessary to provide a full picture of a person's ability? From the tens of thousands of hours of a student's academic career, we can select but a few for careful observation. Through these few hours we need to get as full a picture of the student's ability as possible and thus create an increasingly secure basis for judgment. We can do that by using the other major components of assessment -- observation, judgment, explicit criteria, feedback, and self-assessment -- with a view to shaping a process that makes single assessments complementary and cumulative. Such a process serves the learner by clarifying a pattern that shows the unique highlights and shadows, the fullnesses and gaps of a picture that takes shape gradually with each new line affecting the direction of the next one.

Observation

Because assessment does not depend on predetermined correct answers, it calls forth from teachers their keenest powers of observation. It depends on their ability to set aside any training that has prepared them to quantify and rank, to eliminate possible alternatives, or to limit potential by the "best" they have seen. An effective assessor looks at what is happening behaviorally -- at a student drawing conclusions, for example, whether at a podium or in a paper. Such looking involves attention to parts in precise relationship to each other and to a whole, including emphasis and proportion. It involves adapting an open framework to replace any

tendency one might have to look only for error or to be directed by single-focused expectations.

Such a framework should come from an increasingly expansive understanding of the ability being demonstrated -- an understanding born of an ongoing process of experiencing the ability, reflecting on the experience, and attempting to define it explicitly. Yet the framework must not be limited by that understanding. The framework used by an observer should also be qualitative. It should allow a range of varied expressions and styles -- in fact, assisting the observer to discover them and to describe the contribution of each to overall effectiveness. At the same time, therefore, as the observer verifies identified dimensions of an ability, he or she should explicitly look for dimensions that constitute uniqueness in an individual.

Externality. Assessment requires an assessor to step out of the direct interactive teacher/learner process to observe a developing ability in action. Therefore, it brings in an added outside perspective that helps to extend understanding of student ability. The perspective of externality might come from criteria established department-wide, or from assessing done by someone other than the teacher, or from college-wide assessments that calls forth the integrating of content and/or skills from more than one course.

Even in regular classroom assessments, teachers need to establish a measure of distance to assure that a new judgment is made on the basis of criteria applied to a specific situation, rather than one limited to past perceptions of the student. Otherwise they have no guarantee that their observation and judgment make a fresh addition to the accumulated understanding they have of a student's ability.

In self-assessment, the achieving of some distance is even more of a challenge. The struggle to stand outside of one's own performance to get a different view of it is essentially what makes learning to assess oneself so long and complex a process. Practice -- in looking at records of one's own performance and in general refining of

one's ability to observe and judge according to criteria -- make self-assessment more attainable.

Judgment and Explicit Criteria

The experience of faculty as expert judges of student ability is another important reason for placing them at the center of any educational assessment process. Even faculty who have never verbalized their standards and who might use a norm-referenced framework to report their judgment work out of an understanding of what they expect in student performance. Assessment requires them to articulate that understanding in explicit and public statements of criteria. By doing so, faculty refine their own understanding of expected abilities, clarify for their colleagues the basis of their judgment, and enable students to understand what performance is required.

Explicit criteria provide a major means of getting a picture of an ability, for they serve as indicators of that ability as seen in performance. Thus they are one of the components of assessment that distinguish it as learning. The picture sketched by criteria needs to enable the assessor to judge the presence of an ability. It also needs to be clear enough for the beginning learner to infer a performance from it.

Research on Criteria

Perhaps the most persistent question about explicit criteria is that of determining how specific they should be. Our research at Alverno College suggests that context is an important determining element and that, within the context of individual student learning, the developmental level of the student is significant.

Beginning students. We find that students at the start need very explicit criteria. They are trying to figure out "what they're supposed to do" and, in effect, they use the criteria to plot a performance. They use criteria as a recipe or set of directions. Initial results from the longitudinal study conducted by the Alverno Office of Research and Evaluation support this impression. They indicate that

students begin with the perception that criteria are directions for what and how much to learn and that competences are directions for what to do. While they see explicit directions as "picky," they see broader directions as "vague."

After a semester or two, students begin to cluster the criteria they had formerly seen as discrete and to test these clusters as abilities. For example, they begin to realize that in most cases one cannot just make inferences but must support them with data. But they are often still focusing on a task to be performed. As one student explains, "I'm beginning to realize that I can do well on certain things -- that's confidence I never really had before. I can sit back and say, 'Hey, I really did a good job setting up the transitions between my paragraphs and getting across my ideas on paper!'"

Gradually these students begin to see combined abilities contributing to consistent effectiveness in their performance. One student described her experience this way: "I used to have a hard time assessing how well I was doing. I asked all my friends how they did on their work to see how I compared to the rest of the class. Now I have a goal set in my mind about how well I should be going, and I think I assess my school work now more by my own expectations."

Alverno research shows that as students develop, they begin to see learning as a process, with the criteria providing a picture of given abilities and a framework for self-assessment. Criteria at this level need to be general, yet specific enough to allow students a range of skill in relating criteria to each other, to performance, and to ability.

Advanced students. Students at a later educational stage have begun to create their own picture of an ability. One student put it this way: "Being competent means achieving your goals, achieving not only their criteria, but your criteria. You have your criteria about what you learn. You apply what you learn in your life and your work and see how it comes into play. I think being competent is achieving not only their criteria, but your own personal criteria."

Incomplete as students' pictures might be, specified criteria serve to remind learners what they have not yet internalized or to supplement what they have. Having developed a range of abilities to call on in varied situations, students should be able, given a context, to infer the kind of performance elicited, call upon the required abilities, and infer criteria. According to Alverno research, the most advanced students do begin to internalize the need for criteria; they see criteria as part of self-assessment and use them in a flexible way to guide their learning. Criteria therefore can be more integrated at this level and, when communicated to the student, serve as familiar comparative bases rather than complete surprises.

Interpreted thus, criteria constitute the primary tool of the assessor -- especially if the assessees are expected to learn from the experience and to become assessors of their own performance.

Sequential Administration

If, as we have suggested, assessment can serve learners best when they can carry from one assessment situation to the next a developing picture of their abilities, then assessments need to have connections. Some of those connections learners can make for themselves when faculty identify what is to be assessed, what criteria will be used to judge it, and how well it has been done. But once learners know how well they have done thus far and have an idea of how they might improve, they need opportunities to demonstrate their improvement.

Within a course, therefore, formative assessments need to build on each other in a way that is clear to the student. And summative assessments need to build on the formative and on each other. In fact, if an institution expects of students some outcomes that transcend courses -- and colleges all do, both in relation to the major and general education, then faculty need to provide sequenced, external assessments to give students opportunities to integrate the knowledge and abilities they have demonstrated in discrete courses. In effect, in order to address the student's

ongoing, overall development as a learner, faculty need to extend assessment in random courses to assessment across the curriculum. And that assessing needs to be developmental as well as reinforcing.

Integrated external assessments involving the context of a simulated or actual work place or of a citizen or professional role can keep carrying learning closer to the other settings in which learners primarily use or will be using their abilities. Such assessments also take into consideration the holistic nature of an ability.

Feedback

For assessment as learning, feedback becomes a critical component. It is the teachable moment, the opportunity for change. It takes the elements of assessment discussed thus far and turns them into learning. It can be seen as both a resource and an event. As resource, it is information provided by the assessor and in some cases by the assessment itself which presents a profile of how the learner in action meets criteria of effectiveness. As event, feedback is the time when the learner and assessor "sit down beside each other" and direct their attention to the strengths and weaknesses of the learner's performance.

"Sitting down" can mean that the student and faculty member have a face to face interaction or that a course instructor gives feedback to the entire class and to small groups within the class. It can also take the form of a well-worded sentence written from the faculty member to the learner. Whatever the form, feedback interprets performance as judged by criteria, thus extending the picture of a student's developed ability. It makes this picture available and revealing to both partners in the assessment process.

When feedback is seen as an opportunity to learn rather than the indicator of rank in class or percentage of items correct, then it points to uniquenesses and illuminates the reasons for weaker aspects of the learner's performance. It reveals to the learner the sources of strengths. It suggests where to aim to more fully

develop an ability. By reinforcing the learner's understanding of what he or she knows, it motivates further development. In this later sense, feedback is also a look forward and a time to redirect efforts and make plans to practice nuances of the ability being developed.

Using feedback-as-event, as the teachable moment, also calls for considerable creativity on the part of the teacher and systematic study of the most effective feedback strategies on the part of the profession. Educators are, we think, only at the early stages of developing the art of giving feedback. Although there is much to explore, there are some insights and aids already available; here are a few.

Research on Feedback

Most teachers have found that to be effective, feedback should be timely, informative, explicit, focused on what can be changed, and generally positive in nature. Process-oriented theories of motivation -- such as Expectancy Theory, Operant Learning Theory, and Goal Setting Theory -- support these guidelines and suggest others. Most theories, for example, note the importance of providing the learner with knowledge of results. Yet these and other theories do not clearly specify the conditions under which these guidelines will be most effective. How explicit should feedback be? Should negative feedback be given, and if so, when? What level and amount of information constitutes optimal knowledge of results?

Experience in giving feedback at Alverno suggests that one way to deal with these questions is to study the developmental stages of learners in relation to their use of feedback. While knowledge of "stages" is still incomplete, what the faculty do know and report here has been helpful in working with students.

Beginning Students. Beginning students prefer specific, concrete feedback. They focus on aspects of their performance as if these aspects were isolated, discrete, and unrelated elements. Feedback that is positive, specific, and concrete helps at this

stage, but is most effective if it assists the learner to see the relationships among the discrete elements of performance.

Another characteristic of beginning students (whether they are first year students or beginning a new course of studies sometime later) is that they often let emotional responses hinder their insight. The instructional strategy to give as much positive, specific feedback as possible in earlier assessments. For less successful elements of the performance, instructors also try to develop ways to give feedback that points out why the student ran into difficulty and what concrete steps can be taken in order to improve. Care in these matters is especially important with students for whom knowledge of multiple weaknesses might tend to be overwhelming.

For instance, in a first-semester humanities course at Alverno, students who have written essays on the pros or cons of an aesthetic issue receive feedback on their analytic and writing abilities that concentrates on the positive. Faculty point out such things as how the discussion takes account of the selected audience or where the writer made clear relationships among the key arguments. However, they also point out at least one area that needs further development, as in the following example:

> You show awareness of the author's use of symbolism. Where I think you could improve is in reflecting on the meaning and significance of those symbols. One of the characteristics of a symbol is that it points to some larger idea. You need to be more explicit in identifying those broader areas.

Then, in order to further assure learning, a specific instructor might ask the student to review samples of the work of previous students who had effectively clarified the significance of specific symbols. These samples might be in a reserve file in the department or library. Or they might be called up on a personal computer. Whatever the mode, such feedback challenges learners to move beyond their present ability and, by exposing them to a range of peer examples, gives them some idea of how to do it.

More Advanced Students. As students develop the ability to use feedback as new learning, they take a more objective stance on their own behavior. They seek out

evaluation of their work. They want feedback that helps sort out patterns and relationships among varied abilities and discipline contexts. Consequently, feedback to advanced-level students places less emphasis on elements effectively demonstrated and more emphasis on the learner's performance in relation to past work and to the nuances of the underlying ability.

One strategy Alverno faculty use for giving this type of feedback is to make the expected outcomes of the major as the organizing principle. For instance, history faculty have identified three major outcomes that each graduating major must demonstrate. One of these is the ability to articulate, integrate, and employ methods of history to create a coherent understanding of her own and other cultural heritages. Work submitted during the senior history seminar is assessed in light of this ability, and feedback to students indicates in what way and to what degree each one is demonstrating it. Consequently, feedback to advanced students aims to sketch an increasingly holistic profile of the learner as history major.

Throughout the academic program, faculty at Alverno aim to give feedback geared to the developmental level of the individual. Increasing adeptness at using feedback contributes to the learners' developing ability to assess themselves. Faculty find that even the most successful students find paths of development they have not explored. They begin to develop an experiential definition and thus a potential habit of ongoing learning.

Self-Assessment

As we have indicated earlier in this paper, we see the ability to self-assess, to appraise one's own performance, as the culmination of the learning process: it places within the student's grasp tools and strategies for developing autonomy as a learner. On another level, we see it as an essential component of the assessment process and an important part of each individual assessment. At Alverno, faculty observe that students gradually recognize the importance of the ability to self-assess. It is not

unusual for students to articulate that recognition: "I know what I know and what I don't know. I know what I can do and what my weaknesses are. I think all these assessments, the feedback I've received and the self-assessments have helped me 'take charge' of my own development."

Assisting learners to develop the ability to self-assess is a multi-dimensional process. It means teaching them to observe themselves in action. It requires students to develop the habit of asking what these observations mean about their own behavior and the underlying array of expectations, knowledges, and abilities that these behaviors represent. It asks students to make judgments about the effectiveness of their behavior in reference to a set of standards or criteria rather than making comparisons to the work of peers. It asks them to make distinctions between external and internal sources of their behavior. Finally, developing complex self-assessment ability involves learners in finding more effective, yet distinct, models of performance that can serve as behavioral alternatives for future development.

Such are the goals of the self-assessment process. To nourish the ability is, of course, a difficult task. True, there is a growing body of literature in psychology on the "self" in various research contexts from self-perception and self-efficacy to social information processing. Yet, there are few insights for the educator who seeks to design effective approaches to building this capability. A synthesis of these studies with at least some tentative insights for educators would be a helpful next step. However, for another source of information, one can turn to what students actually say about their self-assessments. Their statements can provide an idea of what they are able to do at different points in their academic program.

Research on Self-Assessment

In interviews with students as they graduate and reviews of statements made by students as part of their assessments, Alverno researchers have identified patterns that characterize the kind of self-assessment beginning, intermediate, and advanced

students are able to do. These patterns have suggested to faculty how to design self-assessment formats that can be helpful at each stage.

Beginning students, for instance, have trouble relating specific, concrete evidence of aspects of their performance to the judgments they make about their work. They tend to make global judgments and, when asked, write down very brief statements, with one eye on what the assessor might think of them. A typical example is a first-year student's comments on a five-page paper she wrote for a communication course:

> It seemed as though I was beating the subject to death. This might be very boring to you.

Alverno faculty have found that having students watch video tapes of their performance in a group setting or reread an essay using a diagnostic framework are effective approaches in teaching beginning learners to observe performance and self-assessment. To help them take more initiative, faculty frequently provide assessments followed by checklists or sets of questions that structure the students' self-assessment in a detailed, systematic manner and help elicit information about their performance that is positive and goal-directed rather than self-deprecating.

Intermediate Students. As learners begin to internalize models of effective performance in different areas of study, they tend to focus on comparing their performance to these models. They then give criteria a special importance in providing an overall picture of the ability they are to demonstrate. And, faculty find that whether students are writing essays in a literature course, analyzing historical documents, or investigating natural phenomena, they are able to look at themselves from a more holistic perspective. They also begin to describe themselves as having a "sense" of whether their own performance in a given situation is essentially competent or not. The following example, from a self-assessment of a speech analyzing two points of view on the media, indicates this more holistic perspective as well as the student's awareness of what might work better:

> I tried to show relationships but in certain areas I think I went overboard and started to explain what was going on in the parable, for example. It's obvious that in each account the people want to be protected from other people -- the media in real life and each other in the parable. I don't think I made this relationship very clear. Perhaps I should have concentrated on two good relationships and expanded on these. On the whole, though, I'm comfortable with what I did.

When students can take a more holistic perspective, they are ready to use self-assessment formats that can assist them to examine their work across a variety of assessments, thus helping them observe patterns and motivate future efforts. One particularly effective method is one in which learners are asked to self-assess on the basis on a course's major outcomes as indicated in the syllabus.

Advanced Students. As learners advance beyond a basic understanding of effectiveness in each area of study, they give evidence of internalizing standards of self-assessment and begin to rely on their own standards for determining the actions they might take. For instance, in a log kept on an off-campus field project for her major area of study, one student wrote:

> I am usually acting on or reacting to directives I must carry out for my supervisor. Usually, I play the role of sustainer. This time I changed my frame of mind and approach to the task, a task I ordinarily wouldn't prefer doing. . . I decided to view it as an opportunity, to prove that I could go the extra mile.

Similarly, in a videotape made during the senior year, a psychology and music education major noted how she used criteria as a reference point but also developed more challenging standards for herself:

> Once I know what the criteria are for the assessment, I set my own standards over and against these. I meet the explicit criteria but try to surpass them with my own. This is when I feel free to be most creative, even if I make a fool out of myself!

Consequently, a number of self-assessments which are designed into senior-level assessments ask the learner to set forth the criteria she would use to evaluate her work. In this way the student is challenged to take more responsibility for evaluating her own performance and is provided an opportunity to expand her conception of what effective performance might mean in the situation.

Uncovering other self-assessment behaviors that characterize students at developing stages during their college years is part of the ongoing research at Alverno. It enables faculty to further develop their theory out of their experience with students and thus to create better designs for self-assessment.

The ability to self-assess, as we see it, assists learners to know what they can do and not attribute their competence, once developed, to luck or to others. It helps them focus on what they experience and pull out what they have learned. In this sense, self-assessment can enable graduates to become their own best teacher and critic as they carry this ability to settings beyond the classroom.

III. DEVELOPING AN ASSESSMENT: GUIDELINES FOR TEACHERS

How does one go about developing an assessment of the kind that we have been describing? And who is the "one" to develop it? How can one do so? It is easy to say, "Take the elements identified above and create a process to embody them." Practitioners know that it is not easy to do. Add to these elements the specific content of a course, discipline, or general education program, the particular level of student to be assessed, and the creative intelligence of a teacher, and the process of assessment design becomes even more complex.

In effect, assessment design explicitly addresses the complexity a teacher is actually dealing with whenever designing a learning experience or system. In order to make the process accessible, therefore, we will deal with it in a practical, inductive fashion, working through the elements as any teacher might, and translate them into a design for developing assessment.

Who is the "one" who develops assessment? In defining assessment as an educational process we have stressed that it includes not only a specific evaluative event, but also the ongoing relationship between teacher and student, and the even more cumulative sense of a student's overall development across the curriculum. Therefore we believe that assessment is a responsibility shared by individual teachers

and a college or university as a whole. We will deal, then, with the design of assessment from two perspectives -- that of the individual teacher thinking it through, then that of the larger curriculum.

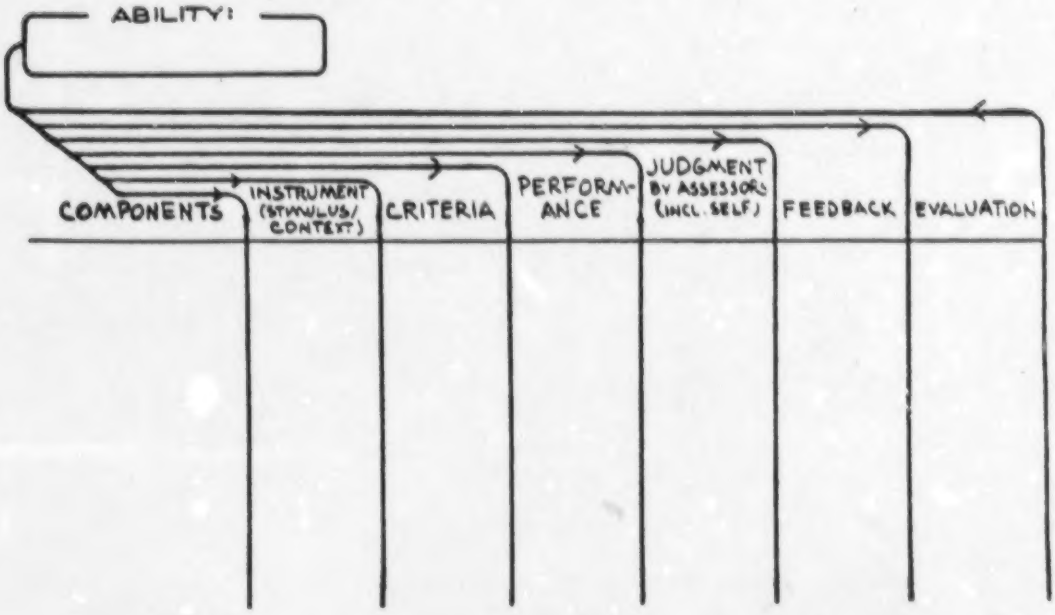
Designing Individual Classroom Assessments

Let us imagine any teacher ready to design an assessment, and thinking aloud: I take as a basic working assumption that my aim is to sample my students' abilities and to provide multiple opportunities for that sampling. I will plan, therefore, to assess my students in an ongoing way, building class experiences as evidence culminating in more formal assessment. I know that the hardest part of my work is to make explicit my criteria for effective performance, and I know that the effectiveness of those criteria depends on how I build in strategies for feedback and self-assessment.

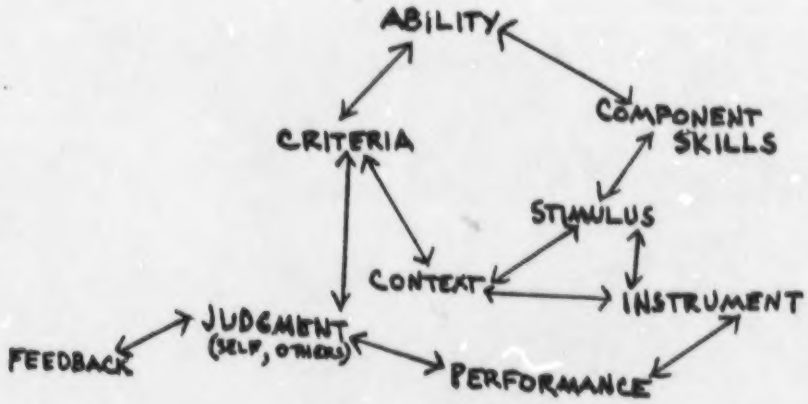
The method that I follow is not a rigid series of steps, but it does follow a logical pattern in relation to the elements of assessment already outlined in this paper. I begin by determining the outcome I expect, the ability I want my students to demonstrate. At some point I have to determine a stimulus and context, and designs for feedback and self-assessment. Beyond any individual assessment, I consider how each assessment experience relates to the ongoing development of the student, especially in relation to other assessments in my course.

Each assessment that I design is part of a larger pattern. I remember what my students have demonstrated in the past; I anticipate what they will be learning in the future. My process, then, includes attention to the development of the student as an individual learner as well as to that student's participation as a member of my particular class. When I design a specific assessment, I try to ensure attention to these several considerations by incorporating all the elements that make up a successful assessment.

At Alverno, a generalized model describes the flow of the process, assuring the inclusion of crucial elements and feeding back into an evaluation of each aspect:



I may not always work with these elements in the same order, but I include them all each time I design an assessment. In use, the model is spread out and rearranged. For my design I can begin at any part, but in doing so, I inevitably set up a series of connections:



In the discussion that follows my starting point is determining the ability to be demonstrated; but wherever I begin, connections between elements lead to the other steps in the process.

1. DETERMINE A SPECIFIC ABILITY OR EXPECTED OUTCOME. A major assumption underlying assessment, as distinct from traditional testing, is that learning -- and by extension, assessment of that learning -- should be designed to foster the growth of student abilities in significant areas beyond the acquisition of knowledge. As I work with students, I teach and assess a variety of abilities ranging from interpreting data to analyzing works for perceivable organizing principles. It is these that I must identify, integrating them with the content of my discipline. In designing an assessment, I examine the overall goals of my course to determine which of them, and with what degree of complexity, I can assess at a given time.

In an introductory fiction course, for example, one of my course goals is for a student to "show understanding of the way in which readers make meaning in literature by analyzing literary elements in relationship with each other."

Such a goal embodies an attempt to focus on the content of my discipline as more than the specific "texts" to be studied -- as not only a knowledge base, but a complex mix of "facts," texts, history, theoretical approaches, analytic frameworks, concepts, interpretations -- and more. My goal, then, deals with "content" as it defines the study of literature. I further particularize this goal through class assignments and assessment when I assign specific stories for consideration.

Breaking Open An Ability into Components

2. IDENTIFYING COMPONENT ABILITIES. Because the ability is a complex one, I need to break it open into component abilities. That step moves me toward the criteria to be used in judging it. In this case it means that my assessment of it does not occur in only one event. Instead, I consider how students can develop this

ability throughout my course, and I plan how to assess for component skills in relation to specific stories which students are reading.

One specific component skill of this overall goal is the ability to identify and discuss literary elements (such as plot, character, tone, style, and so on) with understanding of their use by different writers. This is actually a preliminary step to being able to show how these elements work in relationship to each other as a reader develops a theory of meaning. I assess the preliminary steps so that I can assist students with difficulties they might have at this level before they attempt to demonstrate the more difficult goal.

The complexity or difficulty of the specific stories I assign also varies. As students develop ability to analyze for literary elements in stories that have more accessible style or structure, I can assign more complex works to assess their analytic ability.

In effect, broader goals need to be broken open and spread on a continuum of development. For the beginning student, I set more specific skills to be developed. In an advanced literature course I would not assess students specifically for their ability to use the vocabulary of the discipline, although I might presume this ability and in fact use it as a criterion for assessing a broader goal.

To determine an expected outcome for an assessment, then, a teacher needs to state an ability in relation to the learning content, the course context, the developmental level of the student, the chronology of the assessment event.

3. SELECT OR DESIGN A STIMULUS AND CONTEXT. Although there are those educators who still bristle at a word like "stimulus," it usefully describes that part of an assessment that elicits a student's performance. A stimulus might be a question asked, or an essay presented for analysis, or a problem posed, or an event experienced. It might be a simple request for a choice of answers, or it might be a complex situation in which the possibilities for response are numerous.

Whether I choose a stimulus first and create a context for it, or begin with a context and then find an appropriate stimulus, I have several underlying questions. How will I narrow the content to a concrete situation? Will I be assigning specific texts or events, or problems? a process or product or both? will I ask the student to choose the specific content? how will I limit the choices the student has to make? What do I want students to do with the content to show the ability? write paragraphs? draw diagrams? outline answers? solve problems? By what circumstances will I define limitations? to what audience? for what purpose? with texts available to them? with time constraints? alone or with others? What will prompt them to do it? How will I motivate my students to demonstrate all that they've mastered?

The Importance of Context

I want my students to be motivated to perform well on my assessment, to see their learning as part of their development as competent individuals. Establishing a realistic context is one way to do this, for it helps to break down some of the artificial barriers between the world of the classroom and other worlds. Most of us, as test designers, provide a stimulus by asking questions, but we rarely provide a context beyond the actual test conditions -- open book, or timed, and so on. We ask students to write an essay, or to choose correct answers, or to solve a problem, or to remember data, without a setting or purpose that would relate their actions to anything else they might do.

A teacher might ask students to compare two authors, or presidents, or chemical compounds, without clarifying details that suggest why such a comparison is worth making. Students may or may not be able to imagine those details. And why should they be able to? Expecting them to do so after distracts them from essentials. If I ask students to make a comparison between two authors, it is my responsibility to provide a context. I do so when I ask them to make the comparison based on specific content and theory, for a specific audience, and for a reason: for example, to show

inexperienced readers how language affects our response or to persuade a literary critic that his or her theory of fiction can be questioned with evidence from the authors being compared.

Consideration of Developmental Levels

For the assessment designer, consideration of developmental levels of the students plays as important a role in shaping stimulus and content as it does in articulating abilities to be assessed.

If I am a psychology teacher, for example, I might be assessing beginning students on their ability to show understanding of several significant theories or advanced students on their ability to apply theories appropriately in actual situations. For the beginning student, a good assessment stimulus with context could be: "Create a detailed outline as a study-guide for other members of the class: what will they need to know to master the theories?" The difference here from a question that asks "Explain the major components of these three psychological theories" might seem small, but it is significant. Though the context is simulated, it gives students a realistic purpose and audience to assist in selecting components of the theories and shaping an answer. It provides a framework that focuses the comparison and leads students to the task at hand. It relieves some of the time-absorbing activity of blindly determining a context for themselves -- which is irrelevant to this situation and often misdirected. It relieves others of writing contextless prose made up of generalizations written out of nowhere to no one. At best, it enables them to see their learning as part of their development as competent individuals, able -- and motivated -- to perform.

From the advanced student, I expect a behavior beyond understanding psychological theories, so I provide a different kind of stimulus and context. Perhaps I will ask each student in a group to take on the role of a particular theorist and then pose specific situations to discuss from the unique position of that theorist and in

dialogue with the others. Not only does this stimulus provide the opportunity for an interesting discussion, it also calls forth synthesis and application that carry understanding of the theories far beyond basic knowledge. The discussion mode lessens the students' control over the direction of thought, so that they can show whether they have sufficient understanding of the theories to apply them to whatever situation arises.

Consideration of Mode

In designing a stimulus and context my major focus remains that of developing a situation that will offer the student the best possible chance to show the particular ability I am assessing for. If I am concerned with students' ability to make relationships between literary elements, for example, or to design a nutritional plan for a specific kind of patient, I must devise a "mode" that emphasizes just these abilities. Rather than have students write essays, I might ask the literature students to construct a diagram or map of relationships within a particular story, and the nursing students might be asked to chart a plan. In such cases, students can demonstrate their ability to show relationships in a work of fiction or a nutritional plan without letting other abilities like clear writing distract them.

Choosing a stimulus involves creating a leading question or situation, providing a setting and format for student response in ways appropriate to the specific outcome desired. It also involves recognizing that because a stimulus both shapes and is shaped by the integration of ability and content, my decisions about these elements and their connections can improve by becoming conscious ones.

4. DEVELOP CRITERIA. Whether one sees criteria as the standards by which one judges student performance or as the behavioral indicators by which one fills in their picture of a given ability, the process of developing them means inferring them from performances as experienced and remembered. In designing as assessment, a teacher will have in mind an "ideal performance." Though perhaps not consciously spelled out,

it is part of the inspiration for the need for assessment. As teachers, we have the kind of idea of what we want our students to accomplish that determines what we expect them to do.

My job as assessment designer then is to determine criteria by describing that ideal performance, distinguishing essentials and generalizing enough to accommodate varied styles and varied qualitative aspects. What was good about an analysis of a poem as it was done by a literary scholar, or by a student, or by myself? What was successful in a well-presented speech synthesizing several sociological theories? What made a review of a play valuable to a reader? I might consider performances I remember, or I might imagine a successful performance -- another kind of "remembering." What would I want to see in a good analysis or a good review, or a good speech? My imagined idea of a good performance will probably be based on examples I have in my memory.

I can also determine or refine criteria by literally collecting performances. I build up a sense of what I can expect my students to accomplish as I see what students have actually accomplished.

Whether I base my criteria on remembered, imagined, or collected performances, however, I need to get past that point where the ability is not visible because the picture is too fuzzy. I must say enough in defining criteria to allow the student (and myself as teacher) to recognize the ability that is being assessed.

Identifying Criteria for Analytic Ability

When I am assessing for an ability to analyze, for example, certainly an important aspect of the student's performance is finding specific elements and relationships within a work or process, whether it be a historical document, a scientific formula, or philosophical movement. In identifying criteria, I need to take these broad component skills and work to become more precise about what I expect students to do with elements and relationships as evidence of their ability to

analyze. Good analysis I have seen, for instance, surely includes some immediately identifiable elements: making relationships based on frameworks appropriate to the discipline, clarifying relationships in terms of underlying inferences, supporting relationships by verifiable observations, and stating the significance of relationships in terms of the meaning of the work/process or the effect on the analyzer. Although "relationships" is the focus here, I can see that my statements incorporate "elements."

Specific Considerations in Identifying Criteria

Further adapting criteria to a specific instrument or process involves constantly screening them with considerations of the content specified, the level of quality that defines the ability in a given context, and the developmental level of the student. For content, if I want to assess students' understanding of particular aesthetic conceptual frameworks, I specify that: for example, "clarifies understanding of the relationship between art as an aesthetic construct and art as a reflection of life." If students have already sufficiently demonstrated understanding of specific frameworks, I might give them a choice and a chance to practice using ones they are incorporating into their own aesthetic perspective: for example, "clarifies understanding of relationships between selected aesthetic frameworks." If historical background is important to the ability I am assessing, I also include that: for example, "stating the significance of relationships in terms of trends in literary history."

The number of relationships might be specified for the less experienced student ("at least five") or made part of what is to be assessed for the experienced student ("relationships among major literary elements, major aesthetic frameworks, and major trends in literary history"). In addition to the developmental level of the student, the entire assessment process of my course (How much do I assess elsewhere?) guides my

decision. And a simple consideration like time allotted can keep my decisions realistic in relation to context as well as ideal performance.

Guidelines for Effective Criteria

At Alverno, faculty have codified a set of guidelines by which to judge the criteria they articulate:

1. Do the criteria include enough components to provide a picture of the ability (enough to enable one to infer the ability from them)?
2. Are they appropriate to the level of the student?
3. Are they appropriate to other aspects of context like purpose and mode of stimulus?
4. Could they be met in a variety of ways?
5. Could they be used by another judge?
6. Could they be used to self-assess?

These guidelines can assure any teacher that their more intuitive criteria as well as their more analytic ones do not neglect important aspects.

Must I provide carefully spelled-out criteria for each assessment that I design? Our answer is yes, for several reasons. Whether or not I articulate criteria, I continue to use them and to rely on my expert judgment when I assess student performance. We are convinced that taking the student seriously as learner involves making the basis for this expert judgment available and at the same time refineable by articulating it in the form of criteria for each assessment. How fine or full a picture of expected outcomes individual teachers draw by their criteria remains a function of individual experience and commitment. We have found that the traditionally difficult task of designing good tests and "correcting" them cannot be made easier by an assessment framework, but the difficulty can be rewarded by increasingly visible student learning.

5. PROVIDE FOR SELF-ASSESSMENT. As assessment designer aiming to help my students take responsibility for their own development, I include a dimension beyond their demonstration of a given ability: I ask for some evidence of their own evaluation of that demonstration. By already having completed the hardest part -- designing criteria -- I have provided them with the most important tools for self-assessment. All I still need to build in is a time and stimulus/format for self-assessment. I might include an overall question about the performance or a set of detailed questions about specific aspects of the performance. I might make self-assessment a formal part of the instrument or provide for it more informally through directive suggestions or questions.

Again, the key determining factors for my decisions are the level of the student and the context of this specific assessment. Where is the student in the development of his or her ability to self-assess? How does she use criteria? Does she have an internalized set? Does she have at least the start of a picture of her own strengths and weaknesses in regard to what is being assessed so that she can build on it?

Before considering the range of formats for self-assessment mentioned earlier in this paper, I might decide whether to pitch the self-assessment to a more affective or a more cognitive level. Beginning students especially might benefit most from a question that asks which aspects of the assessment they were able to handle with assurance and which they were unsure of. At other times they might best learn from a request to identify where they had a breakthrough in their thinking while they were working on the assessment. For more advanced students I might ask for self-assessment in a more open-ended way, allowing them to supply their own categories for the self-assessment.

In self-assessment, effective use of criteria remains a useful way for students to see an ability as a whole and to work on it in parts. Sometimes students can best use criteria as checklists to get a larger profile of their ability. Sometimes they can best use them individually as take-off points for further understanding of their

ability. In either case, I try to see that my individual assessments make provisions for self-assessment within the student's developing picture of her actual ability in relation to her potential.

6. JUDGE THE PERFORMANCE AND GIVE FEEDBACK. The judging of the performance and the giving of structured feedback constitute major elements of the assessment design. For the student, these may indeed be the most significant, since judgment and feedback are the visible signs of student success and progress, or the lack of it.

In the assessment design process we have been describing, judgment of the performance is a direct application of developing explicit criteria. As an assessor, I make observations of my student's performance and either expressly record examples of the behavior I observe or at least mentally acknowledge them. On the basis of such evidence, I then judge the student's performance as it meets the criteria I have established. In the context of a course, I would also relate the student's performance to overall development of my course goals. When designing an assessment, I should think ahead of how, within the limitations of my time, I can give feedback that will most benefit the students. Most importantly, I need to generate alternatives: written feedback in a checklist with one focused comment? written as a memo? oral on a tape recorder? oral in face-to-face interviews that replace several lecture periods? in combination with peer feedback? Whatever mode I choose, I generate feedback that is distinct from traditional "grading" in that I focus on giving students a description of how they have performed -- on what they have done rather than on comparison or rating with other students. I describe for the student the successes that I find in the performance even as I make suggestions for ways the performance can be improved. A single criterion measures the quality of any feedback I give: does it add to the student's dynamic picture of his or her own ability in a way that motivates further development?

Designing External General Assessments: An Extended Example

Extending assessment beyond the individual classroom to a wider curriculum context involves collaborative, integrating work by a group of designers from a single discipline or from several. The expected outcomes to be identified become departmental or college-wide. Except for those additional factors, however, the design process is the same: determining expected outcomes, breaking each outcome open into component abilities, creating an instrument, and identifying criteria.

At Alverno, faculty have found it productive to design comprehensive assessments collaboratively. These instruments give them an opportunity to keep clarifying what they mean by general education or by specialization in their major. One such assessment, for example, is an integrated one that Alverno students take near the end of their second year. It is designed to give them a picture of how on a given day in a given situation they are able to bring together the abilities they have developed thus far. Faculty see it as another way of looking at each student's achievement in her general education as a whole. By describing this assessment we can illustrate the challenges, but more especially the successes, of a collaborative effort at assessment design.

The assessment was originally designed by a general education committee at our college. The group had set itself the task of selecting an instrument from which faculty could learn something about each student, and each student could learn something that would assist her in planning her upper division work in her major. They could not find an instrument that directly addressed some of the learning goals they had identified -- aesthetic response, for example, or integrating observations and inferences to clarify meaning in a work or process. Therefore they designed their own, gradually working out a validity study with the Alverno Office of Research and Evaluation.

Abilities to Assess. The design process began with the abilities to be demonstrated. An ability like aesthetic response, for instance, they broke open into

workable components like making judgments about the quality of artistic works and defending judgments on the basis of how an artist sustains audience participation. They agreed on these as important aspects they would want to see in the students' performance. After analyzing other expected abilities, they focused on getting some insight into how ready students would be to pull their abilities into integrated operation on a problem. They wanted to see how students would work with problems that, like the ongoing ones of their personal and professional lives, are not separated neatly into steps and do not come as the direct culmination of preparatory learning experience. Therefore the designers decided on a simulation. It was difficult to come up with one that has some public aspect yet is not related to a specific profession. However, once they thought of having students see themselves as a citizen advisory council to a local school board on the question of censorship of books, both their interest and their creativity became self-propelling. They imagined an entire set of tasks involving interaction with parents, teachers, and news reporters, as well as reading material on academic freedom.

Stimulus and Context. Gradually a half-day assessment took shape that requires students, within time constraints, to read background materials, then prepare and deliver an oral presentation; to deal with a deskful of letters, phone messages, and memos by delegating or providing responses; and finally to develop, in collaboration with four other students, a set of guidelines and recommendations for which each must clarify a rationale. As soon as the faculty designers had the scheme sufficiently completed, they could assign the task of writing imagined scenarios to a creative, articulate teaching assistant and save for themselves the crucial task of identifying criteria.

Examples of Criteria. Because the assessment was to focus on the integration of general education outcomes, the form of which already integrated ability and content, the faculty designers decided to aim for integrated statements of criteria. Two examples show the results:

1. Clearly articulates own position on issue (Integrates valuing in decision making, communication)

2. Identifies implications of and rationale for own position, with accurate reference to and interpretation of a conceptual framework of one of the disciplines studied (Integrates content, analytic ability, valuing in decision making)

Assessing and Administering. Since the assessment integrated and transcended course outcomes, faculty decided that student performances would be judged by external teams -- volunteer professionals from the urban community, teaching assistants, and/or rotating faculty. The assessment would be administered during final assessment week in a situation external to any course. The assessors would also give written and oral feedback, self-assessment would be part of the feedback session.

This collaboratively designed assessment has been used successfully for over ten years at Alverno. The results of the assessment provide ongoing diagnostic and summative feedback to general education instructors and to major departments as well as to individual learners.

Assessing a Major in a Discipline. In any institution of higher education -- whether or not it makes a total commitment to assessment -- individual disciplines and departments have their own kind of opportunity. They can become creative by designing unique assessments that address the abilities a student majoring in their field is expected to develop. By having a student act as a member of a simulated civic cultural center for one week, for example, an English faculty can assess more abilities in a student than they can in a ten-question comprehensive examination. Behavioral science departments can do the same by creating a simulated consulting firms or a research or clinical center. Traditionally, music recitals and art exhibits have provided culminating evidence, and celebration, of developed abilities. Other departments can learn from art and music how to build such dimensions into their assessments; at the same time, fine arts areas can extend recitals and exhibits assessments by identifying explicit outcomes and criteria, by adding tasks that elicit

additional abilities, and by providing vehicles for feedback and meaningful self-assessment.

CONCLUSION

Up to now, we have emphasized the individual teacher as assessment designer and judge of performance, and we have emphasized the potential of individual departments. Teachers or departments can singly try any of the strategies we suggest in order to experience advances in student learning. They can set course goals with a clearer focus on the learner, for instance, and organize instruction and assessment around the goals. Or they can be more explicit with students about learning goals and standards by which student performance will be judged. Teachers or departments can adapt other single aspects of the assessment-as-learning process we have described. They can provide learners with class-time practice in the use of goal-related abilities. They can improve their exams and their feedback by relating them more explicitly to learning goals. We believe that any of these strategies can of itself make an immeasurably helpful inroad for a learner into the unmapped territory of his or her development.

Nonetheless, if faculty act alone as person or department, we believe that is not enough. To work for the learner, assessment calls for a strong series of connections: expected outcomes must connect to criteria for performance, to assessment processes, to instructional strategies. On a day-to-day basis, these connections translate into relatedness between what students learn, how they learn, how they will be judged, and what their learning means for their future. In a collegiate institution, we consider the extent of those connections an important measure of the extent to which the environment is organized for learning. We might make some of these connections in a course or program. But the more assessment is at the heart of the institution, the more its power can serve the learner.

Notes and References

1. The authors of this paper are all three faculty members at Alverno College. Georgine Loacker, Ph.D., is Professor of English. She chairs the Assessment Council and the Analysis/Communication Division. Lucy Cromwell, Ph.D., is Associate Professor of English. She coordinates the Analysis Department and directs a FIPSE funded project in critical thinking. Kathleen O'Brien, Ph.D. candidate, is Assistant Professor of Management. She chairs the Problem-Solving/Social Interaction Division and currently heads a project on Leadership Development sponsored by the National Association for Bank Women and the National Executive Service Corp.

Although the authors named are immediately responsible for this paper, we are indebted to the cumulative thinking of all of our colleagues at Alverno, especially other members of the Assessment Council: Zita Allen, Kathleen Bultman, Margaret Earley, Joyce Fey, George Gurria, Patricia Jensen, Wendell Kringen, Marcia Mentkowski, Glen Rogers, Judeen Schulte, Judith Stanley, Marilyn Thanos, Christine Trimberger, and Allen Wutzdorff.

2. For further concrete examples of actual assessments, see Assessment at Alverno College by Alverno College Faculty (Milwaukee: Alverno Productions, 1985, revised edition). Other sources are G. Loacker, L. Cromwell, J. Fey, and D. Rutherford, Analysis and Communication at Alverno: An Approach to Critical Thinking (Milwaukee: Alverno Productions, 1984) and M. Earley, M. Mentkowski, and J. Schafer, Valuing at Alverno: The Valuing Process in Liberal Education (Milwaukee: Alverno Productions, 1980).

For further examples of student responses, see the above publications and M. Mentkowski and A. Doherty, Careering After College: Establishing the Validity of Abilities Learned in College for Later Careering and Professional Performance. Final

Report to the National Institute of Education: Overview and Summary (Milwaukee: Alverno Productions, 1984, c1983).

A complete list of publications is available from: the Alverno Institute, Alverno College, 3401 South 39 Street, Milwaukee, WI 53215.

Alverno organizes a day each semester for colleagues from other institutions to visit the campus for an overview of the educational program and to confer with faculty, administrators, and students regarding areas of special interest. For information on "A Day at Alverno," write the Alverno Institute, Alverno College, 3401 South 39 Street, Milwaukee, WI 53215.

3. There are numerous publications on the assessment center method in business. For a helpful overview see George C. Thornton III and William C. Byham, Assessment Centers and Managerial Performance (New York: Academic Press, 1982); and Joseph L. Moses and William C. Byham, eds., Applying the Assessment Center Method (New York: Pergamon Press, 1977). The former has an extensive bibliography.

4. For an interpretation of the use of assessment in the National Assessment of Educational Progress see any of the reports published as part of that project (Educational Testing Service, Princeton, New Jersey 08541). For an interpretation of the use of assessment by the Council for Adult and Experiential Learning (CAEL), see any CAEL publications (CAEL Office, Columbia, Maryland 21044).

5. Particularly helpful articles or chapters on shifting trends in testing and Assessment are Robert Glaser, "A Research Agenda for Cognitive Psychology and Psychometrics," American Psychologist 36 (September 1981): 923-936; David C. McClelland, "Testing for Competence Rather Than for 'Intelligence'," American Psychologist 28 (January 1973): 1-14; Warren W. Willingham, "New Methods and

Directions in Achievement Measurement," New Directions for Testing an d Measurement: Measuring Achievement: Progress Over a Decade, no. 5 (San Francisco: Jossey-Bass, 1980).

6. National reports that emphasize the importance of assessment in higher education include Association of American Colleges, Integrity In the College Curriculum: A Report to the Academic Community (Washington, DC: Association of American Colleges, 1985); Study Group on the Conditions of Excellence In American Higher Education, Involvement In Learning: Realizing the Potential of American Higher Education, (Washington, DC: National Institute of Education, 1984); American Association for Higher Education, Current Issues In Higher Education (Washington, DC: American Association for Higher Education, 1979).

7. For a detailed picture of assessment in Great Britain, see John Heywood, Assessment in Higher Education (London: John Wiley & Sons, 1977). Heywood's work includes a comprehensive bibliography.